

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号
特開2001-184840
(P2001-184840A)

(43) 公開日 平成13年7月6日 (2001.7.6)

(51) Int.Cl. ⁷	識別記号	F I	テーマコード*(参考)
G 1 1 B 27/031		G 1 1 B 20/10	G 5 C 0 2 3
G 1 0 L 19/00		H 0 4 N 5/262	5 D 0 4 4
G 1 1 B 20/10		G 1 1 B 27/02	A 5 D 0 4 5
H 0 4 N 5/262		G 1 0 L 9/18	M 5 D 1 1 0
			9 A 0 0 1
審査請求 未請求 請求項の数6 OL (全 9 頁)			

(21) 出願番号 特願平11-362519

(22) 出願日 平成11年12月21日 (1999. 12. 21)

(71) 出願人 000002185

ソニー株式会社
東京都品川区北品川6丁目7番35号

(72) 発明者 横田 哲平

東京都品川区北品川6丁目7番35号 ソニ
ー株式会社内

(72) 発明者 木原 信之

東京都品川区北品川6丁目7番35号 ソニ
ー株式会社内

(74) 代理人 100086841

弁理士 脇 篤夫 (外1名)

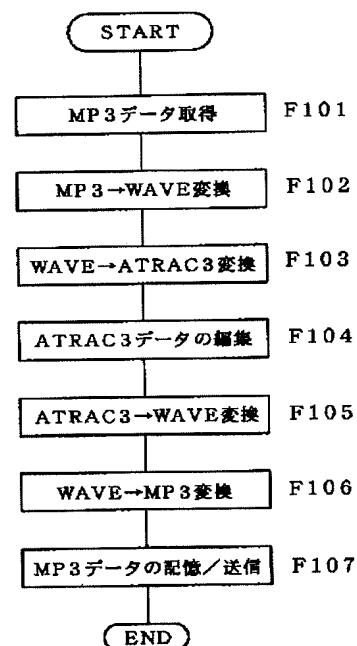
最終頁に続く

(54) 【発明の名称】 編集装置、編集方法

(57) 【要約】

【課題】 MP3などの可変長圧縮方式のオーディオデータに関して編集処理を簡単なものとする。

【解決手段】 可変長の圧縮方式による第1のデータ形態のオーディオデータ (MP3) を、固定長の圧縮方式による第2のデータ形態のオーディオデータ (ATRAC3) に変換して、その第2のデータ形態のオーディオデータについて編集処理を行ない、その後第1のデータ形態のオーディオデータに戻す。



【特許請求の範囲】

【請求項1】 可変長の圧縮方式による第1のデータ形態のオーディオデータを、固定長の圧縮方式による第2のデータ形態のオーディオデータに変換する変換手段と、

前記第2のデータ形態のオーディオデータについて編集処理を行うことのできる編集手段と、

前記編集手段で編集された前記第2のデータ形態のオーディオデータを、前記第1のデータ形態のオーディオデータに変換する変換手段と、

を備えたことを特徴とする編集装置。

【請求項2】 前記第1のデータ形態は、MPEGオーディオレイヤ3であることを特徴とする請求項1に記載の編集装置。

【請求項3】 前記第2のデータ形態は、ATRAC3であることを特徴とする請求項1に記載の編集装置。

【請求項4】 可変長の圧縮方式による第1のデータ形態のオーディオデータを、固定長の圧縮方式による第2のデータ形態のオーディオデータに変換し、
変換された前記第2のデータ形態のオーディオデータについて編集処理を行ない、
編集された前記第2のデータ形態のオーディオデータを、前記第1のデータ形態のオーディオデータに変換することを特徴とする編集方法。

【請求項5】 前記第1のデータ形態は、MPEGオーディオレイヤ3であることを特徴とする請求項4に記載の編集方法。

【請求項6】 前記第2のデータ形態は、ATRAC3であることを特徴とする請求項4に記載の編集方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、例えばMPEGオーディオレイヤ3のように可変長の圧縮方式によるデータ形態のオーディオデータについて好適な編集装置、編集方法に関するものである。

【0002】

【従来の技術】MPEGオーディオレイヤ3（以下、MP3という）と呼ばれるデータ形式により音楽ファイルを記録再生する技術が知られている。このMP3は、例えば16ビット量子化でサンプリング周波数44.1kHz（又は48kHz等）の、PCMリニアオーディオデータ（以下、WAVEデータという）に比べて、音質劣化はさほどない状態で、データ量を約1/11に圧縮することのできるデータ形式である。また、オーディオデータに関する圧縮データ形態としては、ミニディスク（MD）システムで用いられているATRAC方式（Adaptive Transform Acoustic Coding）から発展したATRAC3方式も知られている。

【0003】

【発明が解決しようとする課題】ところでMP3は、主

に音楽配信を目的として開発されたデータ形式であり、特にデータの編集や、記録媒体への書込単位などはさほど考慮されていない。またオーディオデータの1単位となるオーディオユニットについても、1152サンプルのデータが、モードにより必ずしも固定長に圧縮されるわけではない。つまり可変長のデータが許容されている。

【0004】図9はMP3のデータ構造の一例を示している。MP3では、図示するように各フレーム（#1、#2・・・）の先頭に同期信号s yと付加情報a d dが配され、それに続いて主情報（DT1、DT2・・・）、つまりオーディオデータとしての情報が配される。各主情報の先頭には、9ビットの主情報開始位置情報が配され、主情報の開始位置が記述される。主情報のデコードは、その主情報開始位置情報を参照して行われる。

【0005】図9（a）の例では、主情報DT1は比較的小さいデータ量とされており、このためフレーム#1内で完了している。また主情報DT2も小さいデータ量とされており、これもフレーム#1内で完了している。そして、主情報DT3が大容量の場合、主情報DT3はフレーム#1における主情報DT2の後ろから、フレーム#2、#3に渡って配されている状態となっている。主情報DT4は、フレーム#4に配されている。このようなデータ構造にある時に、例えば主情報DT2を削除するとすると、後続するすべての主情報（DT3、DT4）が前ずめに書き換えられることになる。つまり図9（b）のようにデータの書換が行われる。このような処理は、比較的情報量の多いオーディオデータの場合、最悪数10MByteにわたる書換が必要となることもあり、1つの主情報を削除することだけで、その処理を行うソフトウェア、又はハードウェアの負担が非常に大きなものとなっている。

【0006】また、例えば1つの主情報（1つの曲）を、或るポイントで分割して2つの曲に分けるといような編集も、フレームの切れ目と音の編集単位との相関がとれないため、編集のポイントがそのつど変わることになり、編集処理はかなり面倒な処理となる。即ちMP3の44.1kHzのモードでは、1152/44100秒がフレームの単位となっているが、この中から編集ポイントを探し出さなければならない。もちろん上述の削除の場合の曲の切れ目としての編集ポイントを探すことも同様である。さらに2つの曲を1つの曲に連結するという編集の場合も、曲の切れ目を探し出さなければならないため、同様である。さらに、DES等の8の倍数に関係したダミーデータを付加して考えると、MP3の編集処理には一層面倒な処理が要求される。

【0007】このようにMP3形式のオーディオデータは、主情報（曲）の削除や分割などの編集処理には適しているものとはいえない。その一方で、高音質を維持し

て十分に圧縮されてデータサイズとされるMP3は、非常に有用なデータ形式とされていることで、一般に広く利用されているため、ユーザーにとってはMP3で記録したオーディオデータを、より容易に編集したいという要望がある。

【0008】

【課題を解決するための手段】本発明はこのような事情に依拠して、MP3のように可変長の圧縮方式のオーディオデータを容易に編集できるようにした編集装置及び編集方法を提供することを目的とする。

【0009】このため本発明では、可変長の圧縮方式による第1のデータ形態のオーディオデータを、固定長の圧縮方式による第2のデータ形態のオーディオデータに変換する変換手段と、第2のデータ形態のオーディオデータについて編集処理を行うことのできる編集手段と、編集手段で編集された第2のデータ形態のオーディオデータを、第1のデータ形態のオーディオデータに変換する変換手段とを備えるようにする。例えば第1のデータ形態は、MPEGオーディオレイヤ3(MP3)、第2のデータ形態は、ATRAC3であるとする。

【0010】固定長の圧縮方式による第2のデータ形態、例えばATRAC3のオーディオデータは、編集処理が容易である。そこで、第1のデータ形態のデータを第2のデータ形態に変換して編集を行い、その後第1のデータ形態に戻すようにすれば、結果として、第1のデータ形態のオーディオデータの編集が容易に実行できるものとなる。

【0011】

【発明の実施の形態】以下、図1～図8を用いて、本発明の実施の形態の編集装置、編集方法について述べる。この実施の形態では、編集装置としてパーソナルコンピュータを用いた例を挙げる。また第1のデータ形態はMP3、第2のデータ形態はATRAC3であるとする。説明は次の順序で行う。

1. ATRAC3のデータ構造
2. 編集装置となるパーソナルコンピュータの構成
3. 編集処理
4. ATRAC3の編集

【0012】1. ATRAC3のデータ構造

図1にATRAC3のデータ構造を示す。ATRAC3では、図示するように各フレーム(#1、#2・・・)の先頭にヘッダHDと付加情報addが配され、それに続いて主情報(DT1、DT2・・・)、つまりオーディオデータとしての情報が配される。

【0013】各主情報は固定長とされる。フレーム単位は、1024/44100秒である。図1では説明上の例として、各主情報DT1、DT2・・・がそれぞれ1フレームに相当する容量として記してあるが、主情報、即ち1曲のデータ量は、フレーム単位の整数倍となる。そしてフレームの切れ目は音の情報についても切れ目と

なるため、編集処理はフレーム単位で可能となる。即ち図9で説明したMP3の場合のように、フレーム内で曲の削除や分割、或いは結合のための編集ポイントを探し出すということは不要となる。

【0014】2. 編集装置となるパーソナルコンピュータの構成

【0015】編集装置となる本例のパーソナルコンピュータ1の構成を図2に示す。パーソナルコンピュータ1は、CPU11、RAM12、ROM13、ハードディスクドライブ(HDD)14等が設けられており、CPU11により全体の動作制御が行われる。RAM12は、CPU11がプログラムを実行する上でのデータ領域、タスク領域が確保されたり、例えばHDD14等に格納されているアプリケーションプログラム等が展開される部位である。またROM13はパーソナルコンピュータとしての動作プログラムやその他必要な情報が記憶されている。HDD14は、大容量の記憶媒体として、各種のプログラム、データファイルが格納される。本例の場合、外部から取り込んだMP3形式のオーディオファイルもHDD14に格納される。各ブロック間でのファイルデータや制御データの授受はバスBを介して行われる。

【0016】入力装置2はユーザーの操作装置としての各種機器、例えばキーボードやマウス等を示すものであり、入力装置2からの操作入力情報は、入力インターフェース15を介してCPU11に取り込まれる。表示装置3は、CRTや液晶ディスプレイによるモニタ表示装置のことであり、CPU11の指示に基づいて表示ドライブ部17が映像信号を供給することで、モニタ表示が実行される。

【0017】またこの例では、例えばHDD14に格納したMP3形式のデータファイルの再生出力を可能とすること想定しており、このためオーディオデコード部18、D/A変換器19が設けられ、音声信号がスピーカ4に供給できるようにされている。オーディオデコード部18は、CPU11の制御によりHDD14から読み出されたMP3オーディオファイルについてのデコードを行って、WAVEデータ、つまり例えば16ビット量子化、44.1kHzサンプリングのリニアオーディオデータに変換する。そのリニアオーディオデータはD/A変換器19でアナログ信号に変換され、増幅されてスピーカ4に供給されてユーザーに対して出力される。

【0018】通信インターフェース16は、伝送路7、例えばISDN回線などの一般公衆回線或いは何らかの専用回線等と接続してデータ通信を行う部位である。例えばインターネット等を介してのMP3オーディオファイルのダウンロードなどを行う。なお伝送路7は、通信衛星4や各家庭2に設置したパラボラアンテナ5などを利用した衛星通信回線など無線通信回線を想定してもよい。

【0019】外部装置インターフェース20は、コネクタ21を介してパーソナルコンピュータ1と接続された各種の外部装置との間で、データ送受信を行う部位である。図面上ではコネクタ21は1つしか示していないが、複数設けてもよい。また伝送方式は、IEEE1394、USB、IEC958、シリアルポート、パラレルポート、SCSI等、各種考えられる。後述する本例の動作に関していえば、例えばMP3オーディオファイルに関する記録再生を行うことのできるMP3記録再生装置5や、ATRAC3オーディオファイルに関する記録再生を行うことのできるATRAC3記録再生装置がコネクタ21に接続され、これらの機器とパーソナルコンピュータ1の間でMP3オーディオファイルもしくはATRAC3オーディオファイルの伝送を可能とすることが考えられる。

【0020】このような構成によれば、パーソナルコンピュータ1は、通信インターフェース16や外部装置インターフェース20を介して、MP3オーディオファイルを取り込み、HDD14に格納することができる。なお、図示していないが、フロッピーディスクドライブ、CD-ROMドライブ等のデバイスが設けられてもよい。また、アナログオーディオ信号やデジタルオーディオ信号の入力端子や、入力端子から入力されたこれらのオーディオ信号に対するMP3エンコード部を設けることで、入力されたオーディオ信号をMP3オーディオファイルとしてHDD14に格納することができるようにしてもよい。もちろんATRAC3オーディオファイルの記録再生を行うことができるように、ATRAC3エンコーダ、デコーダとしての部位を設けてもよい。

【0021】このパーソナルコンピュータ1がMP3オーディオファイルに関する編集装置として動作するために、例えばソフトウェアによりCPU11に実現される機能ブロック構成を図3に示す。なお、もちろんこのような機能ブロックがハードウェアにより構成されてもよい。

【0022】図3に示すように、まずMP3記録/再生ドライバ31が設けられる。このMP3記録/再生ドライバ31は、HDD14に対するMP3オーディオファイルの書込/読出や、オーディオデコード部18への転送処理などを行う。通信ドライバ32は、通信インターフェース16を介したMP3オーディオファイルの受信又は送信の管理や、外部機器インターフェース20を介したMP3オーディオファイルの受信又は送信の管理を行う。

【0023】MP3/WAVE変換機能33は、MP3オーディオファイルをWAVEデータに変換する処理、及びWAVEデータをMP3オーディオファイルに変換する処理を行う。WAVE/ATRAC3変換機能34はWAVEデータをATRAC3オーディオファイルに変換する処理、及びATRAC3オーディオファイルを

WAVEデータに変換する処理を行う。ATRAC3編集機能35は、ATRAC3オーディオファイルに対する編集処理、例えば曲の削除、分割、連結、曲順移動などの処理を実行する。

【0024】これらの機能ブロックが形成されることで、パーソナルコンピュータ1では、次に説明する編集処理が実行可能となる。

【0025】3. 編集処理

図4によりMP3オーディオファイルに対する編集処理を説明する。図4のステップF101は、編集対象としての特定のMP3オーディオファイルを取得する処理である。例えばHDD14に格納されている或るMP3オーディオファイルについてユーザーが編集処理の実行を求めた場合、CPU11は、MP3記録/再生ドライバ31によりHDD14から1つのMP3オーディオファイルを読み出す。或いは通信ドライバ32により、通信インターフェース16を介して特定のサーバから、もしくは外部機器インターフェース20を介してMP3記録再生装置3から、MP3オーディオファイルを取り込んでよい。

【0026】続いてCPU11はステップF102で、MP3/WAVE変換機能33により、取得したMP3オーディオファイルをWAVEデータに変換する。さらに続いてCPU11はステップF103で、WAVE/ATRAC3変換機能34により、WAVEデータをATRAC3オーディオファイルに変換する。そしてステップF104で、ATRAC3編集機能35により、上記変換処理により得られたATRAC3オーディオファイルに対して、ユーザーの操作に応じた編集処理を実行する。ここで実行される編集例として後に、デバインド（分割）編集、コンバイン（連結）編集について述べる。

【0027】ユーザーの求めた編集処理が完了したら、CPU11はステップF105で、WAVE/ATRAC3変換機能34により、編集後のATRAC3オーディオファイルをWAVEデータに変換する処理を行う。さらにCPU11はステップF106で、MP3/WAVE変換機能33により、WAVEデータをMP3オーディオファイルに変換する処理を行う。そしてステップF107では、このように変換されたMP3オーディオファイルを、編集後のMP3オーディオファイルとして、MP3記録/再生ドライバ31によりHDD14に書き込んだり、或いは通信ドライバ32によりMP3記録再生装置5等に送信することなどを行う。

【0028】以上のように本例では、MP3オーディオファイルについては、一旦ATRAC3オーディオファイルに変換したうえで所要の編集処理を行い、その編集結果としてのATRAC3オーディオファイルをMP3オーディオファイルに戻すことで、MP3オーディオファイルが結果として編集されるようにしている。上述の

ように、ATRAC3データファイルは、編集処理が非常に容易なデータ形式であるため、このようにATRAC3データファイルに変換してから編集処理を行うことで、MP3データファイルの編集が容易に実現できることになる。また一般にパーソナルコンピュータ上のソフトウェアでは、十分に精度の高い浮動小数点演算を自由に行うことができるので、MP3/WAVE変換機能33、WAVE/ATRAC3変換機能34による変換の際の音質劣化は最小限に抑えることができる。

【0029】なお、図4の処理におけるATRAC3データファイル段階（編集直前又は編集直後）において、そのATRAC3データファイルを外部装置インターフェース20を介してATRAC3記録再生装置6に転送し、ATRAC3記録再生装置6における記録媒体に記録させるようにしてもよい。この場合、MP3オーディオファイルをATRAC3記録再生装置6で利用することが可能となる。例えばATRAC3記録再生装置6が、フラッシュメモリを内蔵したメモリカード等に対する記録再生装置であるとする、HDD14に格納されているMP3オーディオファイルをATRAC3オーディオファイルとしてメモリカードに記録させ、そのメモリカードから音楽再生を楽しむことなども可能となる。

【0030】4. ATRAC3の編集
上図4のステップF104で実行されるATRAC3オーディオファイルに対する編集処理例を説明する。

【0031】図5には、1つの曲単位のファイルとしてのATRAC3オーディオファイル（以下、単にオーディオファイルともいう）を示している。このオーディオファイルは、先頭の属性ヘッダと、それに続く実際の暗号化された音楽データとからなる。属性ヘッダは16KB固定長とされる。オーディオファイルの先頭の属性ヘッダは、ヘッダ、1バイトコードの曲名NM1、2バイトコードの曲名NM2、トラックのキー情報等のトラック情報TRKINF、パーツ情報PRTINFと、トラックの付加情報INFとからなる。ヘッダには、総パーツ数、名前の属性、付加情報のサイズの情報等が含まれる。

【0032】このオーディオファイルにおいては、属性ヘッダに対してATRAC3の音楽データが続く。音楽データは、16KBのブロック毎に区切られ、各ブロックの先頭にヘッダが付加されている。

【0033】図6を参照して、曲とATRAC3データファイルの関係について説明する。1曲は、1つのATRAC3データファイル（図5参照）で構成される。ATRAC3データファイルは、ATRAC3により圧縮されたオーディオデータが記録されている。

【0034】1曲は、基本的に1パーツで構成されるが、編集が行われると、複数のパーツから1曲が構成されることがある。パーツとは、録音開始からその停止までの連続した時間内で記録されたデータの単位を意味す

る。1つの曲が複数のパーツで構成される場合、曲内のパーツのつながりは、各曲の属性ヘッダ内のパーツ情報PRTINFで管理される。すなわちパーツ情報PRTINF内に、パーツが持つクラスタの総数、先頭および末尾のクラスタ内の開始サウンドユニット（SUと略記する）の位置、終了SUの位置等が記述される。なおここでいうクラスタとは、図5のブロックに相当する16KByteの単位である。このようなパーツの記述方法を持つことによって、音楽データを編集する際に通常、必要とされる大量の音楽データの移動をなくすることが可能となる。なおブロック単位の編集に限定すれば、同様に音楽データの移動を回避できるが、ブロック単位は、SU単位に比して編集単位が大きすぎる。

【0035】SUは、パーツの最小単位であり、且つATRAC3でオーディオデータを圧縮する時の最小のデータ単位である。44.1kHzのサンプリング周波数で得られた1024サンプル分（ 1024×16 ビット \times 2チャンネル）のオーディオデータを約1/10に圧縮した数百バイトのデータがSUである。1SUは、時間に換算して約23m秒になる。通常は、数千に及ぶSUによって1つのパーツが構成される。1クラスタが42個のSUで構成される場合、1クラスタで約1秒の音を表すことができる。1つの曲を構成するパーツの数は、付加情報サイズに影響される。パーツ数は、1ブロックの中からヘッダや曲名、付加情報データ等を除いた数で決まるために、付加情報が全く無い状態が最大数（645個）のパーツを使用できる条件となる。

【0036】図6は、CD等に記録されたオーディオデータを、2曲の連続したATRAC3オーディオファイルとして或る記録媒体に記録した場合の例を示している。図6（a）に1曲目（オーディオファイル#1）が例えば5つのクラスタ（CL0～CL4）で構成された場合を、また図6（c）に2曲目（オーディオファイル#2）が例えば6つのクラスタ（CL5～CL10）で構成された場合を示している。1曲目と2曲目の曲間では、1クラスタに二つのファイルが混在することが許されない、次のクラスタ（CL5）の最初からオーディオファイル#2が作成される。従って、オーディオファイル#1の終端（1曲目の終端）がクラスタの途中で位置しても、図9（b）に拡大して示すように、そのクラスタの残りの部分には、データ（SU）が存在しないものとされる。第2曲目（オーディオファイル#2）も同様である。そしてこの例の場合は、オーディオファイル#1、#2ともに1パーツで構成される。

【0037】このような形態のATRAC3オーディオファイルに対しては、編集として、デバインド（曲の分割）、コンバイン（曲の連結）、デリート（曲の削除）、ムーブ（曲順移動）などが可能とされる。デバインドは、1つの曲を2つに分割することである。デバインドがされると、総曲数が1つ増加する。デバインドは、一つ

のファイルをファイルシステム上で分割して2つのファイルとする。コンバインは、2つの曲を1つに結合することである。コンバインされると、総曲数が1つ減少する。コンバインは、2つのファイルをファイルシステム上で統合して1つのファイルにする。デリートは、曲を消去することである。消された以降の曲番号が1つ減少する。

【0038】図6に示す二つの曲（オーディオファイル#1、#2）をコンバインした結果を図7に示す。コンバインされたことでオーディオファイル#1、#2は、1つのオーディオファイル#1となり、このオーディオファイル#1は、二つのパーツから形成されるものとなる。上述したように本例ではパーツに関する記述方法があるので、コンバインした結果（図7）において、パーツ1の開始位置、パーツ1の終了位置、パーツ2の開始位置、パーツ2の終了位置をそれぞれSU単位で規定できる。その結果、コンバインした結果のつなぎ目の隙間をつめるために、パーツ2の音楽データを移動する必要がない。

【0039】また、図8は、図6(a)の一つの曲（オーディオファイル#1）をクラスタ2の途中でデバインドした結果を示す。もちろん分割ポイントは図1に示したフレームの切れ目のポイントとなる。デバインドによって、クラスタCL0、CL1およびクラスタCL2の前側からなるデータファイル#1と、クラスタCL2（CL11）の後側とクラスタCL3、CL4とからなるデータファイル#2とが発生する。上述したように、1つのクラスタに二つのファイルが混在することは許されないため、このようにクラスタCL2内の或る位置を分割点とするデバインド編集の場合は、まず、クラスタCL2のデータが、あいている別のクラスタCL11にコピーされる。そしてデータファイル#2においては、クラスタCL11における分割点に相当する位置がスタートポイントとされ、そのクラスタCL11に、クラスタCL3、CL4が続くようにされるものとなる。なお、上述のようにパーツに関する記述方法があるので、デバインドした結果（図8）において、オーディオファイル#2の先頭（クラスタCL11）の空きを詰めるように、データを移動する必要がない。

【0040】以上、実施の形態として本発明の編集装置となるパーソナルコンピュータ1について、その構成及び処理動作を説明してきたが、本発明の編集装置、編集方法としては、多様な変形例が考えられる。例えばパーソナルコンピュータのような形態ではなく、据置型或いは携帯型のオーディオ記録再生装置としての形態で、本

発明の編集装置を実現できる。また、MP3オーディオファイルをATRAC3に変換して編集を行うようにしたが、例えばATRAC、ATRAC2などに変換して編集を行うようにしてもよい。もちろん編集に適した他のデータ形態でもよい。

【0041】

【発明の効果】以上説明したように本発明では、可変長の圧縮方式による第1のデータ形態のオーディオデータを、固定長の圧縮方式による第2のデータ形態のオーディオデータに変換して、その第2のデータ形態のオーディオデータについて編集処理を行ない、その後第1のデータ形態のオーディオデータに戻すようにしたため、例えばMP3などの第1のデータ形態のオーディオデータに関してさほどの処理負担なく編集が可能となるという効果がある。

【図面の簡単な説明】

【図1】本発明の実施の形態でもちいるATRAC3データ構造の説明図である。

【図2】実施の形態の編集装置となるパーソナルコンピュータのブロック図である。

【図3】実施の形態のパーソナルコンピュータのCPU内の機能ブロックの説明図である。

【図4】実施の形態の編集処理のフローチャートである。

【図5】ATRAC3オーディオファイルの構造の説明図である。

【図6】ATRAC3オーディオファイルのパーツ構成の説明図である。

【図7】ATRAC3オーディオファイルのコンバイン編集の説明図である。

【図8】ATRAC3オーディオファイルのデバインド編集の説明図である。

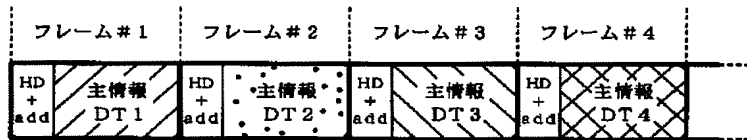
【図9】MP3のデータ構造の説明図である。

【符号の説明】

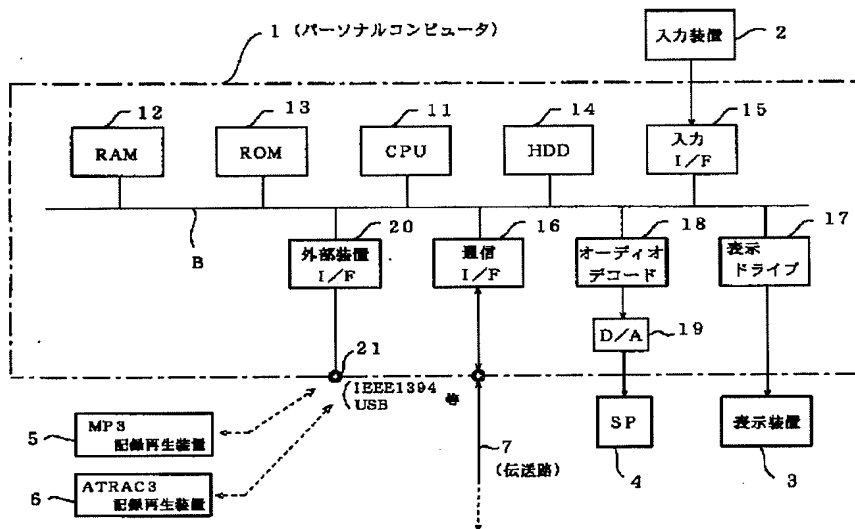
1 パーソナルコンピュータ、2 入力装置、3 表示装置、4 スピーカ、5 MP3記録再生装置、6 ATRAC3記録再生装置、7 伝送路、11 CPU、12 RAM、13 ROM、14 HDD、15 入力インターフェース、16 通信インターフェース、17 表示ドライブ部、18 オーディオデコード部、19 D/A変換器、20 外部装置インターフェース、21 コネクタ、31 MP3記録／再生ドライバ、32 通信ドライバ、33 MP3/WAVE変換機能、34 WAVE/ATRAC3変換機能、35 ATRAC3編集機能

【図1】

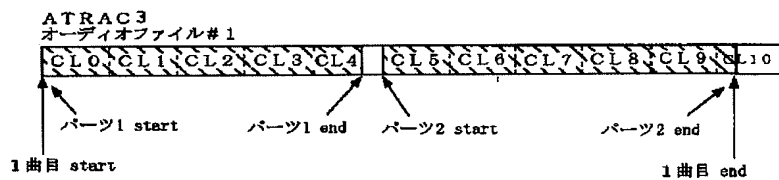
固定長ATRAC3のデータ構造



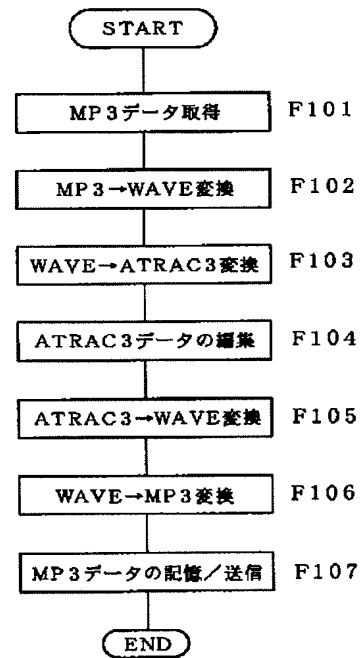
【図2】



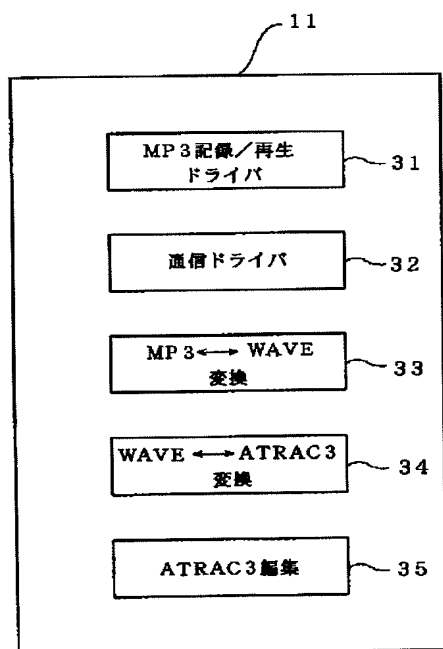
【図7】



【図4】

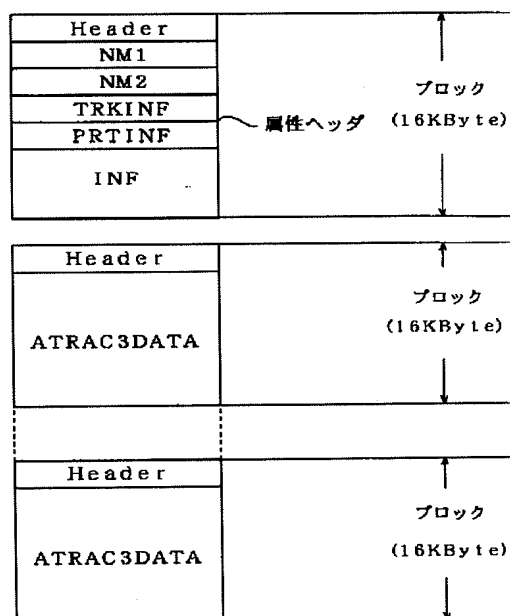


【図3】

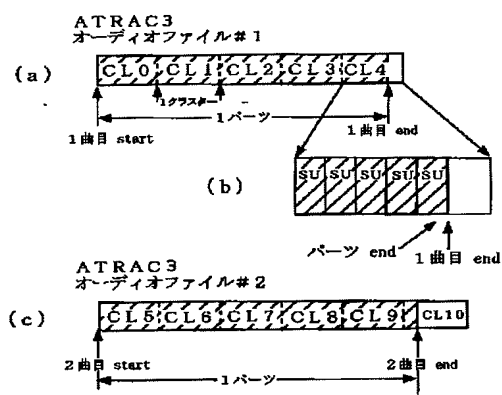


【図5】

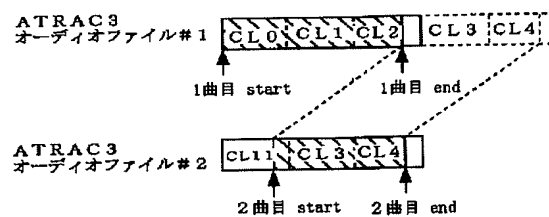
1つのATRAC3オーディオファイル



【図6】

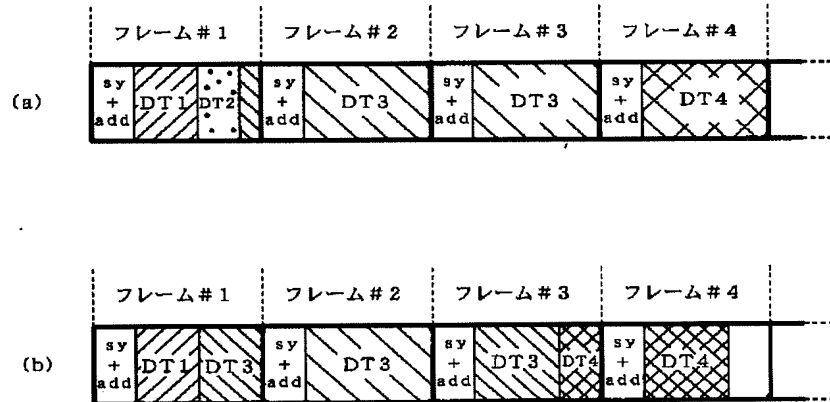


【図8】



【図9】

可変長MP3のデータ構造



フロントページの続き

F ターム(参考) 5C023 CA04 DA03 DA08
 5D044 AB05 BC01 CC05 DE04 DE43
 DE49 GK07 HL02 HL14
 5D045 DA00
 5D110 AA13 AA27 BB02 BB20 CA06
 CA13 CA32 CB04
 9A001 EE02 EE04 HH15 JJ71 KK43
 KK54

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-184840

(43)Date of publication of application : 06.07.2001

(51)Int.Cl. G11B 27/031

G10L 19/00

G11B 20/10

H04N 5/262

(21)Application number : 11-362519 (71)Applicant : SONY CORP

(22)Date of filing : 21.12.1999 (72)Inventor : YOKOTA TEPPEI
KIHARA NOBUYUKI

(54) EDITING DEVICE AND EDITING METHOD

(57)Abstract:

PROBLEM TO BE SOLVED: To simplify editing processing about the audio data of a variable length compression system such as an MP3.

SOLUTION: This device converts the audio data MP3 of a 1st data format by the variable length compression system into audio data ATRAC3 of a 2nd data format by a fixed length compression system, performs editing processing about the audio data of the 2nd data format and subsequently returns the audio data of the 2nd data format to the audio data of the 1st data format.

*** NOTICES ***

JPO and INPIT are not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.*** shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

CLAIMS

[Claim(s)]

[Claim 1]An editing device comprising:

A conversion method which changes audio information of the 1st data form by variable-length compression technology into audio information of the 2nd data form by fixed-length compression technology.

An editing means which can perform editing processing about audio information of said 2nd data form, and a conversion method which changes into audio information of said 1st data form audio information of said 2nd data form edited by said editing means.

[Claim 2]The editing device according to claim 1, wherein said 1st data form is the MPEG audio layer 3.

[Claim 3]The editing device according to claim 1, wherein said 2nd data form is ATRAC3.

[Claim 4]Audio information of the 1st data form by variable-length compression technology is changed into audio information of the 2nd data form by fixed-length compression technology, An editing method performing editing processing about audio information of said 2nd changed data form, and changing audio information of said 2nd edited data form into audio information of said 1st data form.

[Claim 5]The editing method according to claim 4, wherein said 1st data form is the MPEG audio layer 3.

[Claim 6]The editing method according to claim 4, wherein said 2nd data form is ATRAC3.

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[Field of the Invention] This invention relates, for example to an editing device suitable about the audio information of the data form by variable-length compression technology like the MPEG audio layer 3, and an editing method.

[0002]

[Description of the Prior Art] The art which carries out record reproduction of the musical file according to the data format called the MPEG audio layer 3 (henceforth MP3) is known. This MP3 is in the state which tone quality degradation does not have so much compared with the PCM linear audio information (henceforth WAVE data) of 44.1 kHz (or 48 kHz etc.) of sampling frequencies, for example in 16-bit quantization. It is a data format which can compress data volume into the abbreviation 1/11. ATRAC3 method which developed as a compressed data gestalt about audio information from the ATRAC method (Adaptive Transform Acoustic Coding) used by the mini disc (MD) system is also known.

[0003]

[Problem(s) to be Solved by the Invention] By the way, MP3 is the data format mainly developed for the purpose of the music distribution, and in particular edit of data, the write-in unit to a recording medium, etc. are not taken so much into consideration. About the audio unit used as one unit of audio information, the data of 1152 samples is not necessarily compressed into fixed length by the mode. That is, variable-length data is permitted.

[0004] Drawing 9 shows an example of the data structure of MP3. In MP3, the synchronized signal sy and the additional information add are allotted to the head of each frame (#1, #2 ...) so that it may illustrate, and main information (DT1, DT2 ...), i.e., the information as audio information, is allotted following it. 9-bit main information starting position information is allotted to the head of each main information, and the starting position of main information is described. Decoding of main information is performed with reference to the main information starting position information.

[0005] In the example of drawing 9 (a), main information DT1 is made into comparatively small data volume, and, for this reason, it is completed within frame #1. Main information DT2 is made into small data volume, and this is also completed within frame #1. And when main information DT3 is large scale, main information DT3 is in the state where it is allotted over frame #2 and #3 from the back of main information DT2 in frame #1. Main information DT4 is allotted to frame #4. When it is in such a data structure, supposing it deletes main information DT2, for example, all the main information (DT3, DT4) which follows will be rewritten by eye front **. That is, rewriting of data is performed like drawing 9 (b). Since rewriting covering the worst several 10 MByte(s) is needed in the case of audio information with comparatively much amount of information, such processing is only deleting one main information,

and has become what has a very big burden of the software which performs the processing, or hardware.

[0006] Since edit of dividing one main information (one music) on a certain point, and dividing it into two music, for example cannot take correlation with the break of a frame, and the edit unit of a sound, either, it will change the point of edit each time, and editing processing turns into quite troublesome processing. That is, in the 44.1-kHz mode of MP3, although $1152 / 44100$ seconds are a unit of the frame, an edit point must be discovered from this inside. It is also same to, look for the edit point as a break of the music in above-mentioned deletion, of course. Also in edit of connecting two more music with one music, it is the same in order to have to discover the break of music. If the dummy data related to multiples of 8, such as DES, is added and considered, much more troublesome processing will be required of the editing processing of MP3.

[0007] Thus, the audio information of MP3 form cannot say it as a thing suitable for editing processings, such as deletion of main information (music), and division. MP3 which maintains high-quality sound, is fully compressed and is made into data size on the other hand is considered as the very useful data format, and since it is generally used widely, there is a request of liking to edit more easily the audio information recorded by MP3 for a user.

[0008]

[Means for Solving the Problem] An object of this invention is to provide an editing device and an editing method which enabled it to edit audio information of variable-length compression technology easily like MP3 according to such a situation.

[0009] For this reason, a conversion method which changes audio information of the 1st data form by variable-length compression technology into audio information of the 2nd data form by fixed-length compression technology in this invention, It has an editing means which can perform editing processing about audio information of the 2nd data form, and a conversion method which changes into audio information of the 1st data form audio information of the 2nd data form edited by an editing means. For example, suppose that the 1st data form is the MPEG audio layer 3 (MP3), and the 2nd data form is ATRAC3.

[0010] Editing processing is easy for the 2nd data form by fixed-length compression technology, for example, audio information of ATRAC3. Then, if it edits by changing data of the 1st data form into the 2nd data form and is made to return to the 1st data form after that, edit of audio information of the 1st data form can perform easily as a result.

[0011]

[Embodiment of the Invention] Hereafter, the editing device of an embodiment of the invention and an editing method are described using drawing 1 – drawing 8. According to this embodiment, the example which used the personal computer as an editing

device is given. Suppose that the 1st data form is MP3 and the 2nd data form is ATRAC3. Explanation is given in the following order.

1. Edit of composition 3. editing processing 4. ATRAC3 of personal computer used as data structure 2. editing device of ATRAC3 [0012]1. The data structure of ATRAC3 is shown in the data structure diagram 1 of ATRAC3. In ATRAC3, header HD and the additional information add are allotted to the head of each frame (#1, #2 ...) so that it may illustrate, and main information (DT1, DT2 ...), i.e., the information as audio information, is allotted following it.

[0013]Each main information is made into fixed length. Frame units are 1024 / 44100 seconds. drawing 1 -- as the example on explanation -- each main information DT1 and DT2 -- although ... is described as capacity which is equivalent to one frame, respectively, main information, i.e., the data volume of one music, serves as an integral multiple of a frame unit. And since the break of a frame turns into a break also about the information on a sound, the editing processing of it becomes possible per frame. That is, it becomes unnecessary to discover the edit point for deletion of music, division, or combination within a frame like [in the case of MP3 explained by drawing 9].

[0014]2. Composition of personal computer used as editing device [0015]The composition of the personal computer 1 of this example used as an editing device is shown in drawing 2. As for the personal computer 1, CPU11, RAM12, ROM13, and hard disk drive (HDD) 14 grade are provided, and the whole motion control is performed by CPU11. RAM12 is a part where the application program etc. which a data area when CPU11 executes a program, and a task field are secured, or are stored in the HDD14 grade, for example are developed. The operation program and the information required in addition to this as a personal computer are memorized ROM13. The program of the various kinds [14 / HDD] as a mass storage and a data file are stored. In this example, the audio file of the MP3 form incorporated from the exterior is also stored in HDD14. Transfer of the file data during each block or control data is performed via the bus B.

[0016]The input device 2 shows various equipment, for example, the keyboard, a mouse, etc. as a user's manual operating device, and the operational input information from the input device 2 is incorporated into CPU11 via the input interface 15. The display 3 is a monitor display device by CRT or a liquid crystal display, it is that the display driving part 17 supplies a video signal based on directions of CPU11, and a monitor display is performed.

[0017]in this example, the reproducing output of the data file of the MP3 form stored in HDD14, for example is made possible -- thing assumption is carried out, for this reason, the audio decode part 18 and D/A converter 19 are formed, and it enables it to supply an audio signal to the loudspeaker 4 The audio decode part 18 performs decoding about the MP3 audio file read from HDD14 by control of CPU11, and changes it into the linear audio information of WAVE data, i.e., for example, 16-bit

quantization, and a 44.1-kHz sampling. With D/A converter 19, the linear audio information is changed into an analog signal, is amplified, is supplied to the loudspeaker 4, and is outputted to a user.

[0018]The communication interface 16 is a part which connects with ordinary public circuits, such as the transmission line 7, for example, an ISDN circuit etc., or a certain dedicated line, and performs data communications. For example, download of the MP3 audio file through the Internet etc., etc. are performed. The transmission line 7 may assume wireless communications lines, such as a satellite communication line using the parabolic antenna 5 installed in the communications satellite 4 or each home 2, etc.

[0019]The external device interface 20 is a part which performs data transmission and reception among various kinds of external devices connected with the personal computer 1 via the connector 21. On a drawing, although the connector 21 shows only one, they may be provided. [two or more] A transmission system is considered in some numbers IEEE1394, USB, IEC958, a serial port, a parallel port, SCSI, etc. The MP3 recording and reproducing device 5 which can perform record reproduction about an MP3 audio file, for example if it says about the operation of this example mentioned later, ATRAC3 recording and reproducing device which can perform record reproduction about ATRAC3 audio file is connected to the connector 21. It is possible to enable transmission of an MP3 audio file or ATRAC3 audio file between these apparatus and personal computers 1.

[0020]According to such composition, via the communication interface 16 or the external device interface 20, the personal computer 1 can incorporate an MP3 audio file, and can store it in HDD14. Although not illustrated, devices, such as a floppy disk drive and a CD-ROM drive, may be formed. By what is established for the MP3 encode part to these audio signals inputted from the input terminal of an analog audio signal or a digital audio signal, and the input terminal. It may enable it to store in HDD14 by making the inputted audio signal into an MP3 audio file. The part as ATRAC3 encoder and a decoder may be provided so that record reproduction of ATRAC3 audio file can be performed, of course.

[0021]In order that this personal computer 1 may operate as an editing device about an MP3 audio file, the functional block composition realized by CPU11, for example by software is shown in drawing 3. Such a natural functional block may be constituted by hardware.

[0022]As shown in drawing 3, MP3 record / reproduction driver 31 is formed first. This MP3 record / reproduction driver 31 performs writing/read-out of the MP3 audio file to HDD14, transmission processing to the audio decode part 18, etc. The communication driver 32 performs management through the communication interface 16 of reception of an MP3 audio file, or transmission, and management through the external instrument interface 20 of reception of an MP3 audio file, or transmission.

[0023]MP3 / WAVE conversion function 33 performs processing which changes an MP3 audio file into WAVE data, and processing which changes WAVE data into an MP3 audio file. The WAVE/ATRAC3 conversion function 34 performs processing which changes WAVE data into ATRAC3 audio file, and processing which changes ATRAC3 audio file into WAVE data. The ATRAC3 edit function 35 performs processing of the editing processing to ATRAC3 audio file, for example, deletion of music, division, connection, playing order movement, etc.

[0024]By these functional blocks being formed, execution of the editing processing explained below is attained in the personal computer 1.

[0025]3. Editing processing drawing 4 explains the editing processing to an MP3 audio file. Step F101 of drawing 4 is processing which acquires the specific MP3 audio file as an editing object. For example, when a user asks for execution of editing processing about a certain MP3 audio file stored in HDD14, CPU11 reads HDD14 to one MP3 audio file with MP3 record / reproduction driver 31. Or with the communication driver 32, the MP3 recording and reproducing device 3 to an MP3 audio file may also be incorporated from a specific server via the external instrument interface 20 via the communication interface 16.

[0026]Then, CPU11 is Step F102 and changes the acquired MP3 audio file into WAVE data by MP3 / WAVE conversion function 33. Then, CPU11 is Step F103 and changes WAVE data into ATRAC3 audio file by the WAVE/ATRAC3 conversion function 34. And at Step F104, editing processing according to a user's operation is performed by the ATRAC3 edit function 35 to ATRAC3 audio file obtained by the above-mentioned conversion process. Later, DEBAIDO (division) edit and combine (connection) edit are described as an example of edit performed here.

[0027]If the editing processing for which the user asked is completed, CPU11 will be Step F105 and will perform processing which changes ATRAC3 audio file after edit into WAVE data by the WAVE/ATRAC3 conversion function 34. Furthermore, CPU11 is Step F106 and performs processing which changes WAVE data into an MP3 audio file by MP3 / WAVE conversion function 33. And in Step F107, the MP3 audio file changed in this way as an MP3 audio file after edit, It performs writing in HDD14 with MP3 record / reproduction driver 31, or transmitting to MP3 recording and reproducing device 5 grade with the communication driver 32 etc.

[0028]By this example, as mentioned above about an MP3 audio file. After once changing into ATRAC3 audio file, necessary editing processing is performed, and the MP3 audio file is made to be edited by returning ATRAC3 audio file as the edit result to an MP3 audio file as a result. As mentioned above, since editing processing is a very easy data format, after changing ATRAC3 data file into ATRAC3 data file in this way, it is performing editing processing, and edit of an MP3 data file can realize it easily. Generally, by the software on a personal computer, since a sufficiently high-precision floating point arithmetic can be performed freely, the tone quality

degradation in the case of conversion by MP3 / WAVE conversion function 33, and the WAVE/ATRAC3 conversion function 34 can be suppressed to the minimum.

[0029]processing of drawing 4 -- in the ATRAC3 data-file stage (edit direct front stirrup immediately after edit) to kick, The ATRAC3 data file is transmitted to the ATRAC3 recording and reproducing device 6 via the external device interface 20, and it may be made to make it record on the recording medium in the ATRAC3 recording and reproducing device 6. In this case, it becomes possible to use an MP3 audio file with the ATRAC3 recording and reproducing device 6. For example, supposing the ATRAC3 recording and reproducing device 6 is a recording and reproducing device to the memory card etc. which contained the flash memory, It is made to record on a memory card by making into ATRAC3 audio file the MP3 audio file stored in HDD14, and it becomes possible from the memory card to enjoy music reproduction etc.

[0030]4. Explain the example of editing processing over ATRAC3 audio file performed at Step F104 of edit above-mentioned drawing 4 of ATRAC3.

[0031]ATRAC3 audio file (only henceforth an audio file) as a file of one music unit is shown in drawing 5. This audio file consists of a top attribute header and enciphered actual music data following it. An attribute header is made into 16-KB fixed length. The attribute header of the head of an audio file consists of track information TRKINF(s), such as track name NM2 of 1 or 2 bytes of track name NM code of a header and a single byte code, and key information on a track, part information PRTINF, and additional information INF of a track. The information on the total number of multipart forms, the attribute of a name, and the size of additional information, etc. are included in a header.

[0032]In this audio file, the music data of ATRAC3 continues to an attribute header. Music data is divided for 16 KB of every block, and the header is added to the head of each block.

[0033]With reference to drawing 6, the relation between music and ATRAC3 data file is explained. One music comprises one ATRAC3 data file (refer to drawing 5). The audio information into which ATRAC3 data file was compressed by ATRAC3 is recorded.

[0034]Although one music comprises one part fundamentally, when edit is performed, one music may comprise two or more parts. A part means the unit of the data recorded by within a time [which continued from a recording start to the stop]. When one music comprises two or more parts, relation of the part in music is managed by part information PRTINF in the attribute header of each music. That is, the position of the start sound unit (it is written as SU) in the total of the cluster which a part has in part information PRTINF, a head, and the cluster of an end, the position of the end SU, etc. are described. A cluster here is a unit of 16KByte equivalent to the block of drawing 5. By having a describing method of such a part, when editing music data, it becomes possible to usually lose movement of a lot of music data needed. If it limits to

edit of a block unit, movement of music data is avoidable similarly, but as compared with SU unit, the edit unit of a block unit is too large.

[0035]SU is the minimum unit of a part and is the minimum data unit when compressing audio information by ATRAC3. Hundreds of bytes of data which compressed into 10 about 1/10 of audio information for 1024 samples obtained by a 44.1-kHz sampling frequency (1024x16 bits x two channels) is SU. 1SU will be converted into time and will be an about 23-m second. Usually, one part is constituted by SU which attains to thousands. When one cluster comprises 42 SU(s), the sound for about 1 second can be expressed with one cluster. The number of the parts which constitute one music is influenced by additional information size. Since it is decided by the number excluding a header, a track name, additional information data, etc. out of 1 block, the number of multipart forms serves as conditions for which the state where there is no additional information can use the part of the maximum number (645 pieces).

[0036]Drawing 6 shows the example at the time of recording the audio information recorded on CD etc. on a certain recording medium as ATRAC3 continuous audio file of two music. The case where the 2nd (audio file #2) music is constituted from six clusters (CL5-CL10) by drawing 6 (c) when the 1st (audio file #1) music is constituted from five clusters (CL0-CL4) by drawing 6 (a) is shown. Since two files are not allowed to be intermingled in one cluster between the 2nd music with the 1st music, audio file #2 is created from the beginning of the following cluster (CL5). Therefore, as the termination (termination of the 1st music) of audio file #1 is a cluster, even if it is located, as expanded and shown in drawing 9 (b), data (SU) shall not exist in the remaining portion of the cluster. The 2nd music (audio file #2) is the same. And in the case of this example, audio file #1 and #2 comprise one part.

[0037]To ATRAC3 audio file of such a gestalt, DEBAIDO (division of music), a combine (connection of music), deletion (deletion of music), a move (playing order movement), etc. are made possible as edit. DEBAIDO is dividing one music into two. If DEBAIDO is carried out, the total one number of music will increase. DEBAIDO divides one file on a file system, and considers it as two files. A combine is combining two music with one. If a combine is carried out, the total one number of music will decrease. A combine unifies two files on a file system, and carries out them to one file. Deletion is eliminating music. One music number after being erased decreases.

[0038]The result of having carried out the combine of the two music (audio file #1, #2) shown in drawing 6 is shown in drawing 7. By the combine having been carried out, audio file #1 and #2 are set to one audio file #1, and this audio file #1 is formed from two parts. Since there is a describing method about a part in this example as mentioned above, the starting position of the part 1, the end position of the part 1, the starting position of the part 2, and the end position of the part 2 can be specified to the result (drawing 7) of having carried out the combine, per SU, respectively. As a

result, in order to pack the crevice between the knots of the result which carried out the combine, it is not necessary to move the music data of the part 2.

[0039]Drawing 8 shows the result, DEBAIDO [one music (audio file #1) of drawing 6 (a)] in the middle of the cluster 2. Of course, a division point turns into a point of the break of the frame shown in drawing 1. By DEBAIDO, data file #2 which consists of cluster CL3 and CL4 data file #1 which consists of a front side of cluster CL0, CL1, and cluster CL2, and the backside of cluster CL2 (CL11) occurs. Since two files are not allowed to be intermingled in one cluster as mentioned above, in the DEBAIDO edit which makes a division point a certain position in cluster CL2 in this way, it is first copied to cluster CL11 [another] in which the data of cluster CL2 has opened. And the position equivalent to the division point in cluster CL11 is made into the starting point, and it is made for cluster CL3 and CL4 to follow the cluster CL11 in data file #2. Since there is a describing method about a part as mentioned above, it is not necessary to move data so that the opening of the head (cluster CL11) of audio file #2 may be filled in the result (drawing 8), DEBAIDO.

[0040]As mentioned above, about the personal computer 1 which serves as an editing device of this invention as an embodiment, although the composition and processing operation have been explained, various modifications can be considered as the editing device of this invention, and an editing method. For example, the editing device of this invention is realizable with the gestalt as an audio recording and reproducing device not a gestalt like a personal computer but a deferred type, or portable. Although it was made to edit by changing an MP3 audio file into ATRAC3, it may be made to edit by changing into ATRAC, ATRAC2, etc., for example. Other data forms which were suitable for edit, of course may be sufficient.

[0041]

[Effect of the Invention]As explained above, in this invention, the audio information of the 1st data form by variable-length compression technology is changed into the audio information of the 2nd data form by fixed-length compression technology, Since editing processing is performed about the audio information of the 2nd data form and it was made to return to the audio information of the 1st data form after that, it is effective in edit becoming possible without the processing burden like **, for example about the audio information of the 1st data form, such as MP3.

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]

[Drawing 1]It is an explanatory view of the ATRAC3 data structure which have and is by an embodiment of the invention.

[Drawing 2]It is a block diagram of the personal computer used as the editing device of an embodiment.

[Drawing 3]It is an explanatory view of the functional block in CPU of the personal computer of an embodiment.

[Drawing 4]It is a flow chart of the editing processing of an embodiment.

[Drawing 5]It is an explanatory view of the structure of ATRAC3 audio file.

[Drawing 6]It is an explanatory view of the part composition of ATRAC3 audio file.

[Drawing 7]It is an explanatory view of combine edit of ATRAC3 audio file.

[Drawing 8]It is an explanatory view of DEBAIDO edit of ATRAC3 audio file.

[Drawing 9]It is an explanatory view of the data structure of MP3.

[Description of Notations]

1 A personal computer and 2 An input device and 3 A display and 4 Loudspeaker, 5 An MP3 recording and reproducing device, 6 ATRAC3 recording and reproducing device, and 7 Transmission line, 11 CPUs, 12 RAM, 13 ROM, 14 HDD, 15 input interfaces, 16 A communication interface and 17 A display driving part and 18 Audio decode part, 19 A D/A converter and 20 An external device interface and 21 A connector, 31 MP3 record / reproduction driver, 32 communication drivers, 33 MP3 / WAVE conversion function, 34 WAVE/ATRAC3 conversion function, 35 ATRAC3 edit function